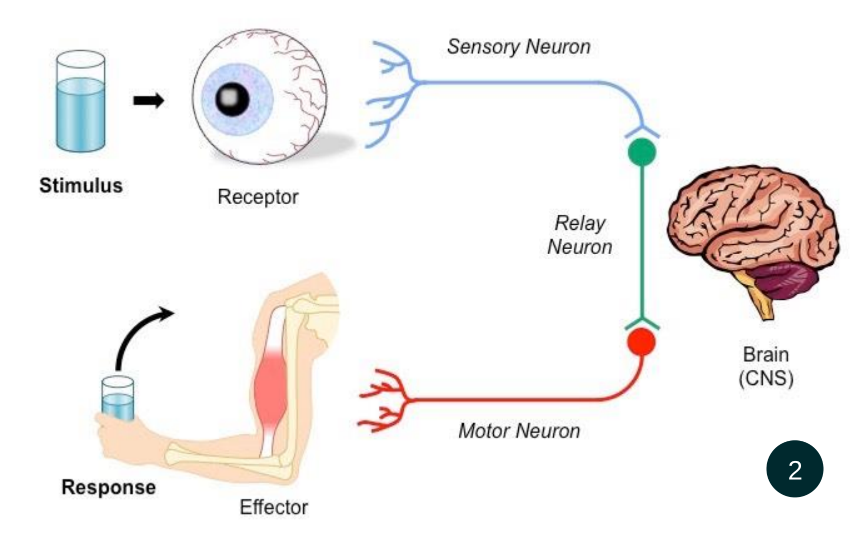
NEURODECODE

UNRAVELING STIMULI FROM NEURAL RESPONSES

Mrigank Maharana
Sept 2023

BACKGROUND

The brain is an intricate network of neurons, each firing electrical signals in response to various stimuli. The pattern of this neural firing, or neural response, can be recorded using various techniques. Each stimulus induces a unique pattern of neural activity. By understanding and decoding these patterns, it's possible to predict the stimulus just by looking at the neural response.



PROBLEM STATEMENT

Develop a system that deciphers or 'decodes' a stimulus (like visual images, sounds, etc.) presented to an organism based on the observed neural responses from the organism's brain.

RESEARCH

Data Collection

Human intracranial EEG (iEEG) recording subdurally or stereotactically with 5 or 10 mm spacing between channels, sampled at 1 kHz or 512 Hz.

Individual data for 10 epileptic patients (mean ± SD [range]: 37 ± 13 [22-69] years of age, 7 males) with channels in frontal and medial temporal lobes. Primary (filtered) and derived (fully preprocessed) iEEG data, and analysis scripts included

Model Selection

a. Model Selection: We opted for two primary models: Logistic Regression and Random Forest. These models were chosen based on their ability to handle high-dimensional data and provide interpretable results.

b. b. Training: Both models were trained on the training dataset

Data Preprocessing

a. Data Extraction: We extracted the neural data and trial summaries from the provided files.
b. Segmenting Neural Data: Neural data, continuous recordings, were segmented based on trial timings to correlate neural responses with individual stimuli.
c. Handling Missing Values: Any columns (features) in the dataset with all missing values were dropped, and other missing values were imputed using the mean strategy.
D. Feature Scaling: Given the potential variance in magnitude across features, we scaled the features using standard scaling.

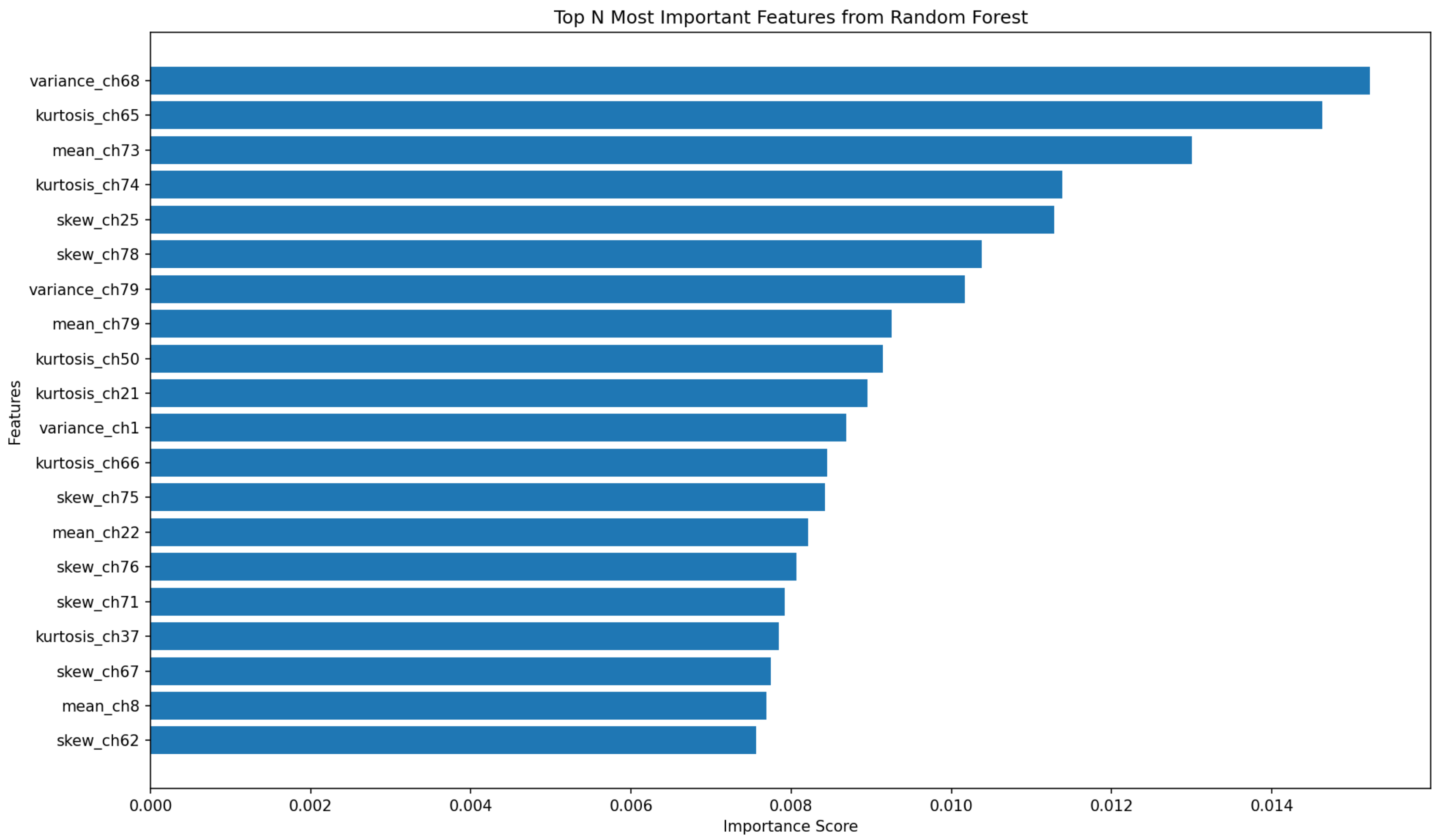
Model Evaluation

The models were evaluated on the test set, and metrics like accuracy and F1-score were used to gauge performance. The Random Forest model also offered feature importance metrics, shedding light on the most influential features in neural decoding.

Feature Engineering

a. Statistical Feature Extraction: From the segmented neural data, we extracted statistical features such as mean, variance, skewness, and kurtosis for each channel.
b. Wavelet Transform: We leveraged the wavelet transform, a tool to decompose signals into different frequency components, to extract additional features from the neural data.
c. Power Spectral Density (PSD): The PSD was computed to extract the power distribution over different frequency components, offering insights into dominant frequencies in the neural response.
D. Data Combination: All extracted features (statistical, wavelet, and PSD) were combined to form a comprehensive feature set for model training.

INITIAL MODEL



INITIAL MODEL

Logistic Regression Accuracy: 70.83%

Metric/Class	Stimulus 1	Stimulus 2	Overall
Precision	0.00	0.67	-
Recall	0.00	1.00	-
F1-Score	0.00	0.80	-
Support	8	16	-
Accuracy	-	-	66.67%
Macro Avg Precision	-	-	0.33
Macro Avg Recall	-	-	0.50
Macro Avg F1-Score	-	-	0.40
Weighted Avg Precision	-	-	0.44
Weighted Avg Recall	-	-	0.67
Weighted Avg F1-Score	-	-	0.53

Note: The dashes ("-") in the table signify that the metric is not applicable for that particular column.

Random Forest Accuracy: 66.67%

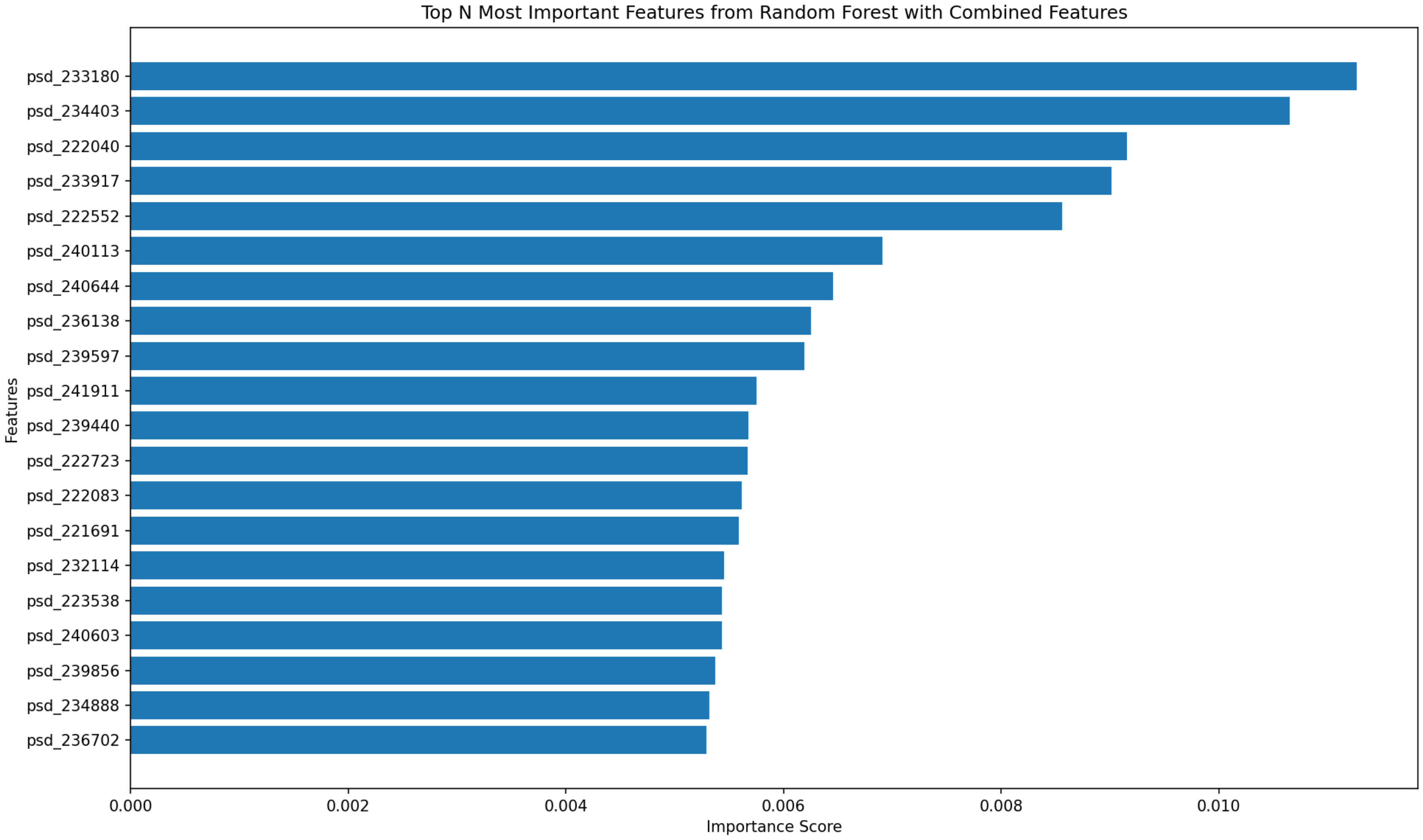
Logistic Regression Results

Measure / Stimulus Type	Stimulus 1	Stimulus 2	Overall
Accuracy	-	-	70.83%
Precision	0.60	0.74	0.69
Recall	0.38	0.88	-
F1-Score	0.46	0.80	-
Support	8	16	24

Macro Average: Precision: 0.67, Recall: 0.62, F1-Score: 0.63, Support: 24

Weighted Average: Precision: 0.69, Recall: 0.71, F1-Score: 0.69, Support: 24

COMBINED MODEL



COMBINED MODEL

Random Forest (with combined features) Accuracy: 83.33%

Metric/Class	Stimulus 1	Stimulus 2	Average
Precision	1.00	0.80	-
Recall	0.50	1.00	-
F1-Score	0.67	0.89	-
Support	8	16	24
Accuracy	-	-	83.33%
Macro Avg	-	-	78%
Weighted Avg	-	-	81%

Note: In the table, "Average" refers to macro and weighted averages where applicable.

Logistic Regression (with combined features) Accuracy: 45.83%

Metric/Class	Stimulus 1	Stimulus 2	Average
Precision	0.22	0.60	-
Recall	0.25	0.56	-
F1-Score	0.24	0.58	-
Support	8	16	24
Accuracy	-	-	45.83%
Macro Avg	-	-	41%
Weighted Avg	-	-	46.5%

Note: In the table, "Average" refers to macro and weighted averages where applicable.

MODEL PERFORMANCE

Initial Model

a. Logistic Regression:
Accuracy: 70.83%

Classification Report:
Stimulus 1: Precision - 0.60, Recall - 0.38, F1-score - 0.46
Stimulus 2: Precision - 0.74, Recall - 0.88, F1-score - 0.80

b. Random Forest:
Accuracy: 66.67%

Classification Report:
Stimulus 1: Precision - 0.00, Recall - 0.00, F1-score - 0.00 (Note: This indicates that the model struggled with this category.)
Stimulus 2: Precision - 0.67, Recall - 1.00, F1-score - 0.80

Combined Feature Model

a. Logistic Regression (with combined features):
Accuracy: 45.83%

Classification Report:
Stimulus 1: Precision - 0.22, Recall - 0.25, F1-score - 0.24
Stimulus 2: Precision - 0.60, Recall - 0.56, F1-score - 0.58

b. Random Forest (with combined features):
Accuracy: 83.33%

Classification Report:
Stimulus 1: Precision - 1.00, Recall - 0.50, F1-score - 0.67
Stimulus 2: Precision - 0.80, Recall - 1.00, F1-score - 0.89

Insights

When trained with combined features, the Random Forest model showed a significant improvement, achieving an accuracy of 83.33%.

The models generally better-decoded stimulus 2 compared to Stimulus 1.

While combining features did not benefit the Logistic Regression model, the Random Forest model greatly benefited from the richer feature set.

WHAT DOES IT MEAN?

Initial Findings: Models showed potential in decoding stimulus from neural data, notably better than a random guess. Both models favored Stimulus 2, hinting that its neural signature might be clearer.

Enhanced Feature Analysis: Logistic Regression: Accuracy dropped to 45.83%, indicating potential noise from added features or overfitting. Random Forest: Accuracy jumped to 83.33%. While it perfectly predicted Stimulus 1 when it did, it only detected half of its occurrences. In contrast, it was highly accurate for Stimulus 2.

Deep Dive: Random Forest particularly benefited from the new features, showcasing the value of these features for decoding tasks. The performance gap between the models underlines the significance of choosing the right model for the data.

Concluding Notes: Results confirm the viability of decoding stimulus from neural signals. Success hinges on the choice of features and the model. Advanced features, like wavelet transformations and power spectral density, showed notable promise in improving decoding accuracy. The differences in model performance highlight the intricate nature of neural-stimulus relationships.

CITATIONS/REFERENCES

(Johnson et al., 2018)Johnson (2018);
Intracranial EEG recordings of medial temporal, lateral frontal, and orbitofrontal regions in 10 human adults performing a visuospatial working memory task.

http://dx.doi.org/10.6080/K0VX0DQD

THANK YOU

Mrigank Maharana

https://github.com/mkorkrish/NeuroDecode

https://mk-maharana.web.app/projects.html

http://crcns.org/data-sets/fcx/fcx-2/about-fcx-2